

SEQUENCE LISTING

<110> Altboum, Zeev
Barry, Eileen M.
Levine, Myron M.

University of Maryland

<120> ISOLATION AND CHARACTERIZATION OF THE
CSA OPERON

<130> UOFMD.006A

<150> 60/198,626

<151> 2000-04-20

<160> 40

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 717

<212> DNA

<213> E. coli

<220>

<221> CDS

<222> (1)...(717)

<400> 1

```

atg cat aaa tta ttt tgt tta cta agt tta ctc ata act cca ttt gtt 48
Met His Lys Leu Phe Cys Leu Leu Ser Leu Leu Ile Thr Pro Phe Val
      5              10              15

gca aat gca aac ttt atg ata tat cca ata tca aaa gat tta aag aat 96
Ala Asn Ala Asn Phe Met Ile Tyr Pro Ile Ser Lys Asp Leu Lys Asn
      20              25              30

gga aat agc gag tta att cgt gtt tat tca aaa tca aaa gag ata caa 144
Gly Asn Ser Glu Leu Ile Arg Val Tyr Ser Lys Ser Lys Glu Ile Gln
      35              40              45

tat ata aaa ata tat aca aaa aag att att aat ccc ggc aca act gaa 192
Tyr Ile Lys Ile Tyr Thr Lys Lys Ile Ile Asn Pro Gly Thr Thr Glu
      50              55              60

gaa cat gaa gtt gat atg ccc aat tgg gat ggt ggg ttt gta gtt act 240
Glu His Glu Val Asp Met Pro Asn Trp Asp Gly Gly Phe Val Val Thr
      65              70              75              80

cct caa aaa gtt att ctt cct gca gga ggg agt aaa tca ata cgt tta 288
Pro Gln Lys Val Ile Leu Pro Ala Gly Gly Ser Lys Ser Ile Arg Leu
      85              90              95

act caa ttt aga ata cca aaa aaa gag gaa att tat aga gta tat ttt 336
Thr Gln Phe Arg Ile Pro Lys Lys Glu Glu Ile Tyr Arg Val Tyr Phe
      100             105             110

gag gcg gta aaa cca gat agc aaa gaa aat gta att gat aat aaa aaa 384
Glu Ala Val Lys Pro Asp Ser Lys Glu Asn Val Ile Asp Asn Lys Lys
      115             120             125

```

Volume in drive D:\ is 010420_1412
Directory of D:\

seqlist.txt

89

4/20/01

1 file(s)
Total filesize 89 KB
0 kilobytes free

"F002H0" 4686E860

SEQUENCE LISTING

<110> Altboum, Zeev
Barry, Eileen M.
Levine, Myron M.

University of Maryland

<120> ISOLATION AND CHARACTERIZATION OF THE
CSA OPERON

<130> UOFMD.006A

<150> 60/198,626

<151> 2000-04-20

<160> 40

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 717

<212> DNA

<213> E. coli

<220>

<221> CDS

<222> (1)...(717)

<400> 1

atg	cat	aaa	tta	ttt	tgt	tta	cta	agt	tta	ctc	ata	act	cca	ttt	gtt	48
Met	His	Lys	Leu	Phe	Cys	Leu	Leu	Ser	Leu	Leu	Ile	Thr	Pro	Phe	Val	
1				5					10					15		

gca	aat	gca	aac	ttt	atg	ata	tat	cca	ata	tca	aaa	gat	tta	aag	aat	96
Ala	Asn	Ala	Asn	Phe	Met	Ile	Tyr	Pro	Ile	Ser	Lys	Asp	Leu	Lys	Asn	
			20					25					30			

gga	aat	agc	gag	tta	att	cgt	gtt	tat	tca	aaa	tca	aaa	gag	ata	caa	144
Gly	Asn	Ser	Glu	Leu	Ile	Arg	Val	Tyr	Ser	Lys	Ser	Lys	Glu	Ile	Gln	
		35					40					45				

tat	ata	aaa	ata	tat	aca	aaa	aag	att	att	aat	ccc	ggc	aca	act	gaa	192
Tyr	Ile	Lys	Ile	Tyr	Thr	Lys	Lys	Ile	Ile	Asn	Pro	Gly	Thr	Thr	Glu	
	50					55					60					

gaa	cat	gaa	gtt	gat	atg	ccc	aat	tgg	gat	ggg	ggg	ttt	gta	gtt	act	240
Glu	His	Glu	Val	Asp	Met	Pro	Asn	Trp	Asp	Gly	Gly	Phe	Val	Val	Thr	
	65				70					75					80	

cct	caa	aaa	gtt	att	ctt	cct	gca	gga	ggg	agt	aaa	tca	ata	cgt	tta	288
Pro	Gln	Lys	Val	Ile	Leu	Pro	Ala	Gly	Gly	Ser	Lys	Ser	Ile	Arg	Leu	
				85					90						95	

act caa ttt aga ata cca aaa aaa gag gaa att tat aga gta tat ttt 336
 Thr Gln Phe Arg Ile Pro Lys Lys Glu Glu Ile Tyr Arg Val Tyr Phe
 100 105 110

gag gcg gta aaa cca gat agc aaa gaa aat gta att gat aat aaa aaa 384
 Glu Ala Val Lys Pro Asp Ser Lys Glu Asn Val Ile Asp Asn Lys Lys
 115 120 125

cta aca aca gag cta tct gtt aat ata att tat gcg gct cta atc aga 432
 Leu Thr Thr Glu Leu Ser Val Asn Ile Ile Tyr Ala Ala Leu Ile Arg
 130 135 140

tct tta cca agt gaa caa aac ata tca cta aac att tct aga aat gca 480
 Ser Leu Pro Ser Glu Gln Asn Ile Ser Leu Asn Ile Ser Arg Asn Ala
 145 150 155 160

aga aaa aat ata att att tat aat aat ggg aat gtt aga gca ggt gtt 528
 Arg Lys Asn Ile Ile Ile Tyr Asn Asn Gly Asn Val Arg Ala Gly Val
 165 170 175

aaa gat att tat ttt tgt aag tca tct aat atc gat gat agc tgt gta 576
 Lys Asp Ile Tyr Phe Cys Lys Ser Ser Asn Ile Asp Asp Ser Cys Val
 180 185 190

aaa aaa acg cat aac aag aat ata tat cca gaa aag tca ttt gat acg 624
 Lys Lys Thr His Asn Lys Asn Ile Tyr Pro Glu Lys Ser Phe Asp Thr
 195 200 205

ctg gtt aat aac aat ttt tct tat gtt ttc att aaa tta aac cat gaa 672
 Leu Val Asn Asn Asn Phe Ser Tyr Val Phe Ile Lys Leu Asn His Glu
 210 215 220

gac ata gaa aaa gag caa gga cta ata caa tta aaa gtt cct tga 717
 Asp Ile Glu Lys Glu Gln Gly Leu Ile Gln Leu Lys Val Pro *
 225 230 235

<210> 2
 <211> 238
 <212> PRT
 <213> E. coli

<400> 2
 Met His Lys Leu Phe Cys Leu Leu Ser Leu Leu Ile Thr Pro Phe Val
 1 5 10 15
 Ala Asn Ala Asn Phe Met Ile Tyr Pro Ile Ser Lys Asp Leu Lys Asn
 20 25 30
 Gly Asn Ser Glu Leu Ile Arg Val Tyr Ser Lys Ser Lys Glu Ile Gln
 35 40 45
 Tyr Ile Lys Ile Tyr Thr Lys Lys Ile Ile Asn Pro Gly Thr Thr Glu
 50 55 60
 Glu His Glu Val Asp Met Pro Asn Trp Asp Gly Gly Phe Val Val Thr
 65 70 75 80
 Pro Gln Lys Val Ile Leu Pro Ala Gly Gly Ser Lys Ser Ile Arg Leu

tct	aca	aaa	caa	ctc	ccc	att	act	gta	tca	tat	gga	gga	aag	act	cta	336
Ser	Thr	Lys	Gln	Leu	Pro	Ile	Thr	Val	Ser	Tyr	Gly	Gly	Lys	Thr	Leu	
		100						105					110			
tca	acc	gca	gat	gtg	act	ttt	gaa	cct	gca	gaa	tta	aat	ttt	gga	acg	384
Ser	Thr	Ala	Asp	Val	Thr	Phe	Glu	Pro	Ala	Glu	Leu	Asn	Phe	Gly	Thr	
		115					120					125				
tca	ggt	gta	act	ggt	gta	tct	tct	tcc	caa	gat	tta	gtg	att	ggt	gcg	432
Ser	Gly	Val	Thr	Gly	Val	Ser	Ser	Ser	Gln	Asp	Leu	Val	Ile	Gly	Ala	
		130					135				140					
act	aca	gca	caa	gca	cca	acg	gcg	gga	aat	tat	agt	ggg	gtc	gtt	tct	480
Thr	Thr	Ala	Gln	Ala	Pro	Thr	Ala	Gly	Asn	Tyr	Ser	Gly	Val	Val	Ser	
					150				155						160	
atc	tta	atg	acc	tta	gca	tca	taa									504
Ile	Leu	Met	Thr	Leu	Ala	Ser	*									
				165												

<210> 4
 <211> 167
 <212> PRT
 <213> E. coli

<400> 4
 Met Lys Leu Lys Lys Thr Ile Gly Ala Met Ala Leu Thr Thr Met Phe
 1 5 10 15
 Val Ala Met Ser Ala Ser Ala Val Glu Lys Asn Ile Thr Val Thr Ala
 20 25 30
 Ser Val Asp Pro Thr Ile Asp Ile Leu Gln Ala Asp Gly Ser Ser Leu
 35 40 45
 Pro Thr Ala Val Glu Leu Thr Tyr Ser Pro Ala Ala Ser Arg Phe Glu
 50 55 60
 Asn Tyr Lys Ile Ala Thr Lys Val His Thr Asn Val Ile Asn Lys Asn
 65 70 75 80
 Val Leu Val Lys Leu Val Asn Asp Pro Lys Leu Thr Asn Val Leu Asp
 85 90 95
 Ser Thr Lys Gln Leu Pro Ile Thr Val Ser Tyr Gly Gly Lys Thr Leu
 100 105 110
 Ser Thr Ala Asp Val Thr Phe Glu Pro Ala Glu Leu Asn Phe Gly Thr
 115 120 125
 Ser Gly Val Thr Gly Val Ser Ser Ser Gln Asp Leu Val Ile Gly Ala
 130 135 140
 Thr Thr Ala Gln Ala Pro Thr Ala Gly Asn Tyr Ser Gly Val Val Ser
 145 150 155 160
 Ile Leu Met Thr Leu Ala Ser
 165

<210> 5
 <211> 2604
 <212> DNA
 <213> E. coli

<220>

<221> CDS

<222> (1)...(2604)

<400> 5

atg	aca	aaa	aaa	aat	aca	tta	tat	ata	acg	atc	atc	gca	atg	cta	act	48
Met	Thr	Lys	Lys	Asn	Thr	Leu	Tyr	Ile	Thr	Ile	Ile	Ala	Met	Leu	Thr	
1				5					10					15		
cca	tat	tca	gtt	ttt	tcc	gga	gat	ata	ccc	aac	tct	ttc	cgt	gat	tta	96
Pro	Tyr	Ser	Val	Phe	Ser	Gly	Asp	Ile	Pro	Asn	Ser	Phe	Arg	Asp	Leu	
			20					25					30			
tgg	gga	gaa	caa	gat	gaa	ttt	tat	gaa	gta	aaa	cta	tat	gga	caa	act	144
Trp	Gly	Glu	Gln	Asp	Glu	Phe	Tyr	Glu	Val	Lys	Leu	Tyr	Gly	Gln	Thr	
		35					40					45				
cta	gga	ata	cat	cga	att	aaa	aca	acc	cca	aca	cat	att	aag	ttt	tat	192
Leu	Gly	Ile	His	Arg	Ile	Lys	Thr	Thr	Pro	Thr	His	Ile	Lys	Phe	Tyr	
		50				55					60					
tca	ccc	gaa	agc	att	tta	gat	aaa	ata	aat	gta	aaa	aaa	gaa	aag	gaa	240
Ser	Pro	Glu	Ser	Ile	Leu	Asp	Lys	Ile	Asn	Val	Lys	Lys	Glu	Lys	Glu	
	65				70				75						80	
aag	aaa	ttg	agt	gtt	ttg	ttc	act	aat	tct	ttt	tca	aga	aat	ggc	aat	288
Lys	Lys	Leu	Ser	Val	Leu	Phe	Thr	Asn	Ser	Phe	Ser	Arg	Asn	Gly	Asn	
				85				90						95		
atg	agt	tgt	cag	ggg	aat	gct	act	ata	cag	tat	aac	tgc	aat	tac	att	336
Met	Ser	Cys	Gln	Gly	Asn	Ala	Thr	Ile	Gln	Tyr	Asn	Cys	Asn	Tyr	Ile	
			100					105					110			
aaa	aca	aaa	tca	gta	gat	gtc	atc	gtt	gat	gat	gtt	gat	aat	gtt	gtt	384
Lys	Thr	Lys	Ser	Val	Asp	Val	Ile	Val	Asp	Asp	Val	Asp	Asn	Val	Val	
		115					120					125				
aac	ctt	ttt	ata	ggg	aat	gaa	ttt	ctg	gat	tct	gaa	gca	cac	aat	gat	432
Asn	Leu	Phe	Ile	Gly	Asn	Glu	Phe	Leu	Asp	Ser	Glu	Ala	His	Asn	Asp	
		130				135					140					
gaa	tat	cat	caa	tta	tca	cga	aat	gta	aaa	aaa	gct	ttt	ata	caa	agc	480
Glu	Tyr	His	Gln	Leu	Ser	Arg	Asn	Val	Lys	Lys	Ala	Phe	Ile	Gln	Ser	
	145				150				155						160	
cag	aca	att	aat	gtc	tca	gat	tct	ggg	aag	tat	aaa	agt	ttg	tct	gtt	528
Gln	Thr	Ile	Asn	Val	Ser	Asp	Ser	Gly	Lys	Tyr	Lys	Ser	Leu	Ser	Val	
				165				170						175		
tca	ggg	aat	agc	gcg	ctg	ggg	att	aca	gat	aca	agt	tat	gct	gtc	tta	576
Ser	Gly	Asn	Ser	Ala	Leu	Gly	Ile	Thr	Asp	Thr	Ser	Tyr	Ala	Val	Leu	
			180					185					190			
aat	tgg	tgg	atg	aat	tac	aat	aaa	ttt	aat	ggg	tac	agc	aac	aac	gaa	624

Asn Trp Trp Met Asn Tyr Asn Lys Phe Asn Gly Tyr Ser Asn Asn Glu	
195 200 205	
aga aca atc aat agt ttg tac ttt aga cat gat tta gat aag aga tat	672
Arg Thr Ile Asn Ser Leu Tyr Phe Arg His Asp Leu Asp Lys Arg Tyr	
210 215 220	
tat tat caa ttt gga cga atg gat cgt aca gat ttg tca caa agt att	720
Tyr Tyr Gln Phe Gly Arg Met Asp Arg Thr Asp Leu Ser Gln Ser Ile	
225 230 235 240	
agc ggg aac ttt aat ttt aac tta ctt cct tta ccc gat att gat ggt	768
Ser Gly Asn Phe Asn Phe Asn Leu Leu Pro Leu Pro Asp Ile Asp Gly	
245 250 255	
ata agg aca gga acc aca caa tct tat atc aaa aat aca gat aag ttt	816
Ile Arg Thr Gly Thr Thr Gln Ser Tyr Ile Lys Asn Thr Asp Lys Phe	
260 265 270	
atc gca tcc cct gta act gtt atg tta act aat ttt tcc aga gtg gaa	864
Ile Ala Ser Pro Val Thr Val Met Leu Thr Asn Phe Ser Arg Val Glu	
275 280 285	
gct ttt cgc aat aat caa tta ttg ggc gta tgg tat tta gat tct gga	912
Ala Phe Arg Asn Asn Gln Leu Leu Gly Val Trp Tyr Leu Asp Ser Gly	
290 295 300	
gta aat gaa tta gat aca gct cgt tta cct tat ggt agt tac gat ctt	960
Val Asn Glu Leu Asp Thr Ala Arg Leu Pro Tyr Gly Ser Tyr Asp Leu	
305 310 315 320	
aaa tta aaa att ttt gaa aat act cag tta gtt cgt gaa gaa ata att	1008
Lys Leu Lys Ile Phe Glu Asn Thr Gln Leu Val Arg Glu Glu Ile Ile	
325 330 335	
cct ttt aat aaa ggg aga agt tct att ggt gat atg caa tgg gac gtt	1056
Pro Phe Asn Lys Gly Arg Ser Ser Ile Gly Asp Met Gln Trp Asp Val	
340 345 350	
ttc att cag gga ggg aat att att aat gac aag gat cgt tac ata gaa	1104
Phe Ile Gln Gly Gly Asn Ile Ile Asn Asp Lys Asp Arg Tyr Ile Glu	
355 360 365	
aaa caa aat aat cat aag tca tca gtt aat gct ggg cta cgt tta cca	1152
Lys Gln Asn Asn His Lys Ser Ser Val Asn Ala Gly Leu Arg Leu Pro	
370 375 380	
att acg aaa aat atc tct gtt caa caa gga gca tct gtt ata gat aat	1200
Ile Thr Lys Asn Ile Ser Val Gln Gln Gly Ala Ser Val Ile Asp Asn	
385 390 395 400	
aaa aat tat tat gaa ggg agt ctc aaa tgg aat tcc ggc att ctg tct	1248
Lys Asn Tyr Tyr Glu Gly Ser Leu Lys Trp Asn Ser Gly Ile Leu Ser	
405 410 415	

ggc tca cta aat agt gag ttc agt ttt ctt tgg gga gat aat gca aaa	1296
Gly Ser Leu Asn Ser Glu Phe Ser Phe Leu Trp Gly Asp Asn Ala Lys	
420 425 430	
ggt aat tat caa agt atc tcg tat acc gat gga ttt agt tta tca ttt	1344
Gly Asn Tyr Gln Ser Ile Ser Tyr Thr Asp Gly Phe Ser Leu Ser Phe	
435 440 445	
tat cat aat gat aag cgg gtc gat aat tgt gga aga aat tac aat gct	1392
Tyr His Asn Asp Lys Arg Val Asp Asn Cys Gly Arg Asn Tyr Asn Ala	
450 455 460	
ggt tgg agt gga tgc tac gaa tca tat tcg gca tct tta agt att cct	1440
Gly Trp Ser Gly Cys Tyr Glu Ser Tyr Ser Ala Ser Leu Ser Ile Pro	
465 470 475 480	
tta ttg gga tgg aca agt act ctg gga tat agt gac act tat agt gaa	1488
Leu Leu Gly Trp Thr Ser Thr Leu Gly Tyr Ser Asp Thr Tyr Ser Glu	
485 490 495	
tca gtt tat aaa aac cat att ctt tct gaa tat ggt ttt tat aat caa	1536
Ser Val Tyr Lys Asn His Ile Leu Ser Glu Tyr Gly Phe Tyr Asn Gln	
500 505 510	
aac ata tat aaa ggg aga acc caa aga tgg caa ctg act tcg tcc acc	1584
Asn Ile Tyr Lys Gly Arg Thr Gln Arg Trp Gln Leu Thr Ser Ser Thr	
515 520 525	
tct tta aaa tgg atg gat tat aat ttt atg cca gca att gga ata tat	1632
Ser Leu Lys Trp Met Asp Tyr Asn Phe Met Pro Ala Ile Gly Ile Tyr	
530 535 540	
aac agt gag caa aga caa ctg act gat aaa ggc gga tat ata tct gta	1680
Asn Ser Glu Gln Arg Gln Leu Thr Asp Lys Gly Gly Tyr Ile Ser Val	
545 550 555 560	
act ctc acc cga gcc agc aga gaa aat tca tta aac gca ggg tat tct	1728
Thr Leu Thr Arg Ala Ser Arg Glu Asn Ser Leu Asn Ala Gly Tyr Ser	
565 570 575	
tac aac tat tcc aga gga aag tat tct tct aac gaa tta ttt gtt gat	1776
Tyr Asn Tyr Ser Arg Gly Lys Tyr Ser Ser Asn Glu Leu Phe Val Asp	
580 585 590	
gga tat atg aca tca aca aat aat ggt gac tat cat gag gta aga atg	1824
Gly Tyr Met Thr Ser Thr Asn Asn Gly Asp Tyr His Glu Val Arg Met	
595 600 605	
cgt ttt aat aaa aat aga cat aat gca gaa ggt aga ctt tca ggt cgt	1872
Arg Phe Asn Lys Asn Arg His Asn Ala Glu Gly Arg Leu Ser Gly Arg	
610 615 620	
ata aac aat cga ttt gga gat tta aat ggt tca ttc agc atg aat aaa	1920
Ile Asn Asn Arg Phe Gly Asp Leu Asn Gly Ser Phe Ser Met Asn Lys	
625 630 635 640	

aac aga aac acc aac agt agc aat cat tct ctc act ggt ggt tat aat	1968
Asn Arg Asn Thr Asn Ser Ser Asn His Ser Leu Thr Gly Gly Tyr Asn	
645 650 655	
tcc tca ttt gct ctt aca agt gat gga ttt tac tgg gga gga agt gca	2016
Ser Ser Phe Ala Leu Thr Ser Asp Gly Phe Tyr Trp Gly Gly Ser Ala	
660 665 670	
tct ggt ttg aca aaa cta gct ggc ggt att atc aag gtt aaa tca aac	2064
Ser Gly Leu Thr Lys Leu Ala Gly Gly Ile Ile Lys Val Lys Ser Asn	
675 680 685	
gat act aaa aaa aat ctg gta aaa gtg act ggg gca ttg tac ggt gat	2112
Asp Thr Lys Lys Asn Leu Val Lys Val Thr Gly Ala Leu Tyr Gly Asp	
690 695 700	
tat tcg cta ggg agc aac gat aat gct ttt att cct gta cca gca tta	2160
Tyr Ser Leu Gly Ser Asn Asp Asn Ala Phe Ile Pro Val Pro Ala Leu	
705 710 715 720	
act cca gcc agt tta att att gaa gat aat aat tat ggt gac aag aat	2208
Thr Pro Ala Ser Leu Ile Ile Glu Asp Asn Asn Tyr Gly Asp Lys Asn	
725 730 735	
att tct gta ctt gca cca acg aac aac gat atg ttt ata ttg ccg ggt	2256
Ile Ser Val Leu Ala Pro Thr Asn Asn Asp Met Phe Ile Leu Pro Gly	
740 745 750	
aat gtt tat cct gtt gaa att gaa acc aaa gta agt gtt tct tat att	2304
Asn Val Tyr Pro Val Glu Ile Glu Thr Lys Val Ser Val Ser Tyr Ile	
755 760 765	
ggt aga ggt ttt gac aaa aac ggc acg cca ctt tct ggc gca cat gtt	2352
Gly Arg Gly Phe Asp Lys Asn Gly Thr Pro Leu Ser Gly Ala His Val	
770 775 780	
ttg aat gaa cca cat gtt atc ctg gat gag gac ggt gga ttt tcg ttt	2400
Leu Asn Glu Pro His Val Ile Leu Asp Glu Asp Gly Gly Phe Ser Phe	
785 790 795 800	
gaa tat aca ggt aat gag aaa aca ctt ttt tta tta aag ggc aga act	2448
Glu Tyr Thr Gly Asn Glu Lys Thr Leu Phe Leu Leu Lys Gly Arg Thr	
805 810 815	
att tat aca tgt caa ctg ggg aaa aat aaa gtt cac aaa ggc att gtt	2496
Ile Tyr Thr Cys Gln Leu Gly Lys Asn Lys Val His Lys Gly Ile Val	
820 825 830	
ttc gtc gga gat gtt ata tgt gat gtt aat agc aca agt tcc tta cca	2544
Phe Val Gly Asp Val Ile Cys Asp Val Asn Ser Thr Ser Ser Leu Pro	
835 840 845	
gat gaa ttt gta aag aac cca cgt gtg cag gat ttg ctg gca aag aat	2592
Asp Glu Phe Val Lys Asn Pro Arg Val Gln Asp Leu Leu Ala Lys Asn	

850

855

860

gat aaa gga taa
 Asp Lys Gly *
 865

2604

<210> 6
 <211> 867
 <212> PRT
 <213> E. coli

<400> 6

Met	Thr	Lys	Lys	Asn	Thr	Leu	Tyr	Ile	Thr	Ile	Ile	Ala	Met	Leu	Thr
1				5					10					15	
Pro	Tyr	Ser	Val	Phe	Ser	Gly	Asp	Ile	Pro	Asn	Ser	Phe	Arg	Asp	Leu
			20					25					30		
Trp	Gly	Glu	Gln	Asp	Glu	Phe	Tyr	Glu	Val	Lys	Leu	Tyr	Gly	Gln	Thr
		35					40					45			
Leu	Gly	Ile	His	Arg	Ile	Lys	Thr	Thr	Pro	Thr	His	Ile	Lys	Phe	Tyr
		50				55					60				
Ser	Pro	Glu	Ser	Ile	Leu	Asp	Lys	Ile	Asn	Val	Lys	Lys	Glu	Lys	Glu
65					70					75					80
Lys	Lys	Leu	Ser	Val	Leu	Phe	Thr	Asn	Ser	Phe	Ser	Arg	Asn	Gly	Asn
				85					90					95	
Met	Ser	Cys	Gln	Gly	Asn	Ala	Thr	Ile	Gln	Tyr	Asn	Cys	Asn	Tyr	Ile
			100					105					110		
Lys	Thr	Lys	Ser	Val	Asp	Val	Ile	Val	Asp	Asp	Val	Asp	Asn	Val	Val
		115					120					125			
Asn	Leu	Phe	Ile	Gly	Asn	Glu	Phe	Leu	Asp	Ser	Glu	Ala	His	Asn	Asp
		130				135					140				
Glu	Tyr	His	Gln	Leu	Ser	Arg	Asn	Val	Lys	Lys	Ala	Phe	Ile	Gln	Ser
145					150					155					160
Gln	Thr	Ile	Asn	Val	Ser	Asp	Ser	Gly	Lys	Tyr	Lys	Ser	Leu	Ser	Val
				165					170					175	
Ser	Gly	Asn	Ser	Ala	Leu	Gly	Ile	Thr	Asp	Thr	Ser	Tyr	Ala	Val	Leu
			180					185					190		
Asn	Trp	Trp	Met	Asn	Tyr	Asn	Lys	Phe	Asn	Gly	Tyr	Ser	Asn	Asn	Glu
	195					200						205			
Arg	Thr	Ile	Asn	Ser	Leu	Tyr	Phe	Arg	His	Asp	Leu	Asp	Lys	Arg	Tyr
	210				215						220				
Tyr	Tyr	Gln	Phe	Gly	Arg	Met	Asp	Arg	Thr	Asp	Leu	Ser	Gln	Ser	Ile
225					230					235					240
Ser	Gly	Asn	Phe	Asn	Phe	Asn	Leu	Leu	Pro	Leu	Pro	Asp	Ile	Asp	Gly
				245					250					255	
Ile	Arg	Thr	Gly	Thr	Thr	Gln	Ser	Tyr	Ile	Lys	Asn	Thr	Asp	Lys	Phe
			260					265					270		
Ile	Ala	Ser	Pro	Val	Thr	Val	Met	Leu	Thr	Asn	Phe	Ser	Arg	Val	Glu
		275					280					285			
Ala	Phe	Arg	Asn	Asn	Gln	Leu	Leu	Gly	Val	Trp	Tyr	Leu	Asp	Ser	Gly
	290					295					300				
Val	Asn	Glu	Leu	Asp	Thr	Ala	Arg	Leu	Pro	Tyr	Gly	Ser	Tyr	Asp	Leu
305					310					315					320
Lys	Leu	Lys	Ile	Phe	Glu	Asn	Thr	Gln	Leu	Val	Arg	Glu	Glu	Ile	Ile
				325					330					335	

Pro	Phe	Asn	Lys	Gly	Arg	Ser	Ser	Ile	Gly	Asp	Met	Gln	Trp	Asp	Val
			340					345					350		
Phe	Ile	Gln	Gly	Gly	Asn	Ile	Ile	Asn	Asp	Lys	Asp	Arg	Tyr	Ile	Glu
		355					360					365			
Lys	Gln	Asn	Asn	His	Lys	Ser	Ser	Val	Asn	Ala	Gly	Leu	Arg	Leu	Pro
	370					375					380				
Ile	Thr	Lys	Asn	Ile	Ser	Val	Gln	Gln	Gly	Ala	Ser	Val	Ile	Asp	Asn
385					390					395					400
Lys	Asn	Tyr	Tyr	Glu	Gly	Ser	Leu	Lys	Trp	Asn	Ser	Gly	Ile	Leu	Ser
				405					410					415	
Gly	Ser	Leu	Asn	Ser	Glu	Phe	Ser	Phe	Leu	Trp	Gly	Asp	Asn	Ala	Lys
			420					425					430		
Gly	Asn	Tyr	Gln	Ser	Ile	Ser	Tyr	Thr	Asp	Gly	Phe	Ser	Leu	Ser	Phe
		435					440					445			
Tyr	His	Asn	Asp	Lys	Arg	Val	Asp	Asn	Cys	Gly	Arg	Asn	Tyr	Asn	Ala
	450					455					460				
Gly	Trp	Ser	Gly	Cys	Tyr	Glu	Ser	Tyr	Ser	Ala	Ser	Leu	Ser	Ile	Pro
465					470					475					480
Leu	Leu	Gly	Trp	Thr	Ser	Thr	Leu	Gly	Tyr	Ser	Asp	Thr	Tyr	Ser	Glu
				485					490					495	
Ser	Val	Tyr	Lys	Asn	His	Ile	Leu	Ser	Glu	Tyr	Gly	Phe	Tyr	Asn	Gln
			500					505					510		
Asn	Ile	Tyr	Lys	Gly	Arg	Thr	Gln	Arg	Trp	Gln	Leu	Thr	Ser	Ser	Thr
		515					520					525			
Ser	Leu	Lys	Trp	Met	Asp	Tyr	Asn	Phe	Met	Pro	Ala	Ile	Gly	Ile	Tyr
	530					535					540				
Asn	Ser	Glu	Gln	Arg	Gln	Leu	Thr	Asp	Lys	Gly	Gly	Tyr	Ile	Ser	Val
545					550					555					560
Thr	Leu	Thr	Arg	Ala	Ser	Arg	Glu	Asn	Ser	Leu	Asn	Ala	Gly	Tyr	Ser
				565					570					575	
Tyr	Asn	Tyr	Ser	Arg	Gly	Lys	Tyr	Ser	Ser	Asn	Glu	Leu	Phe	Val	Asp
			580					585					590		
Gly	Tyr	Met	Thr	Ser	Thr	Asn	Asn	Gly	Asp	Tyr	His	Glu	Val	Arg	Met
		595					600					605			
Arg	Phe	Asn	Lys	Asn	Arg	His	Asn	Ala	Glu	Gly	Arg	Leu	Ser	Gly	Arg
	610					615					620				
Ile	Asn	Asn	Arg	Phe	Gly	Asp	Leu	Asn	Gly	Ser	Phe	Ser	Met	Asn	Lys
625					630					635					640
Asn	Arg	Asn	Thr	Asn	Ser	Ser	Asn	His	Ser	Leu	Thr	Gly	Gly	Tyr	Asn
				645					650					655	
Ser	Ser	Phe	Ala	Leu	Thr	Ser	Asp	Gly	Phe	Tyr	Trp	Gly	Gly	Ser	Ala
			660					665					670		
Ser	Gly	Leu	Thr	Lys	Leu	Ala	Gly	Gly	Ile	Ile	Lys	Val	Lys	Ser	Asn
		675					680					685			
Asp	Thr	L													

770 775 780
 Leu Asn Glu Pro His Val Ile Leu Asp Glu Asp Gly Gly Phe Ser Phe
 785 790 795 800
 Glu Tyr Thr Gly Asn Glu Lys Thr Leu Phe Leu Leu Lys Gly Arg Thr
 805 810 815
 Ile Tyr Thr Cys Gln Leu Gly Lys Asn Lys Val His Lys Gly Ile Val
 820 825 830
 Phe Val Gly Asp Val Ile Cys Asp Val Asn Ser Thr Ser Ser Leu Pro
 835 840 845
 Asp Glu Phe Val Lys Asn Pro Arg Val Gln Asp Leu Leu Ala Lys Asn
 850 855 860
 Asp Lys Gly
 865

<210> 7
 <211> 330
 <212> DNA
 <213> E. coli

<220>
 <221> CDS
 <222> (1) ... (330)

<400> 7
 atc agt aag ttg gca gca tca cct gta ttt ctt gaa aga ggg gtg aat 48
 Ile Ser Lys Leu Ala Ala Ser Pro Val Phe Leu Glu Arg Gly Val Asn
 1 5 10 15

 ata tct gta aga ata cag aag caa att tta tca gaa aaa cca tat gtt 96
 Ile Ser Val Arg Ile Gln Lys Gln Ile Leu Ser Glu Lys Pro Tyr Val
 20 25 30

 gca ttc aga ttg aac gga gac ata cta aga cat tta aag gat gca ttg 144
 Ala Phe Arg Leu Asn Gly Asp Ile Leu Arg His Leu Lys Asp Ala Leu
 35 40 45

 atg ata ata tat ggt atg tca aaa ata gat acc aat gat tgt aga aat 192
 Met Ile Ile Tyr Gly Met Ser Lys Ile Asp Thr Asn Asp Cys Arg Asn
 50 55 60

 atg tca agg aaa ata atg aaa aca gaa gtg gat aaa acc tta ctg gat 240
 Met Ser Arg Lys Ile Met Lys Thr Glu Val Asp Lys Thr Leu Leu Asp
 65 70 75 80

 gta tta aaa aat ata aat agc tat gat gac tca gct ttt ata tct aat 288
 Val Leu Lys Asn Ile Asn Ser Tyr Asp Asp Ser Ala Phe Ile Ser Asn
 85 90 95

 ttg ata tat tta att tca aag atc gag aat aat aaa aaa taa 330
 Leu Ile Tyr Leu Ile Ser Lys Ile Glu Asn Asn Lys Lys *
 100 105

<210> 8
 <211> 109
 <212> PRT
 <213> E. coli

<400> 8
 Ile Ser Lys Leu Ala Ala Ser Pro Val Phe Leu Glu Arg Gly Val Asn
 1 5 10 15
 Ile Ser Val Arg Ile Gln Lys Gln Ile Leu Ser Glu Lys Pro Tyr Val
 20 25 30
 Ala Phe Arg Leu Asn Gly Asp Ile Leu Arg His Leu Lys Asp Ala Leu
 35 40 45
 Met Ile Ile Tyr Gly Met Ser Lys Ile Asp Thr Asn Asp Cys Arg Asn
 50 55 60
 Met Ser Arg Lys Ile Met Lys Thr Glu Val Asp Lys Thr Leu Leu Asp
 65 70 75 80
 Val Leu Lys Asn Ile Asn Ser Tyr Asp Asp Ser Ala Phe Ile Ser Asn
 85 90 95
 Leu Ile Tyr Leu Ile Ser Lys Ile Glu Asn Asn Lys Lys
 100 105

<210> 9
 <211> 1086
 <212> DNA
 <213> E. coli

<220>
 <221> CDS
 <222> (1)...(1086)

<400> 9
 atg aat aag att tta ttt att ttt aca ttg ttt ttc tct tca gta ctt 48
 Met Asn Lys Ile Leu Phe Ile Phe Thr Leu Phe Phe Ser Ser Val Leu
 1 5 10 15
 ttt aca ttt gct gta tcg gca gat aaa att ccc gga gat gaa agc ata 96
 Phe Thr Phe Ala Val Ser Ala Asp Lys Ile Pro Gly Asp Glu Ser Ile
 20 25 30
 act aat att ttt ggc ccg cgt gac agg aac gaa tct tcc ccc aaa cat 144
 Thr Asn Ile Phe Gly Pro Arg Asp Arg Asn Glu Ser Ser Pro Lys His
 35 40 45
 aat ata tta aat aac cat att aca gca tac agt gaa agt cat act ctg 192
 Asn Ile Leu Asn Asn His Ile Thr Ala Tyr Ser Glu Ser His Thr Leu
 50 55 60
 tat gat agg atg act ttt tta tgt ttg tct tct cac aat aca ctt aat 240
 Tyr Asp Arg Met Thr Phe Leu Cys Leu Ser Ser His Asn Thr Leu Asn
 65 70 75 80
 gga gca tgt cca acc agt gag aat cct agc agt tca tcg gtc agc ggt 288
 Gly Ala Cys Pro Thr Ser Glu Asn Pro Ser Ser Ser Ser Val Ser Gly
 85 90 95

305 310 315 320

atc agt gtt ccg gtg ttg tgt tgg cct gga cgt ttg caa ttg gat gca 1008
 Ile Ser Val Pro Val Leu Cys Trp Pro Gly Arg Leu Gln Leu Asp Ala
 325 330 335

aaa gtg gaa aat ccc gag gct gga caa tat atg ggt aat att aat gtt 1056
 Lys Val Glu Asn Pro Glu Ala Gly Gln Tyr Met Gly Asn Ile Asn Val
 340 345 350

act ttc aca cca agt agt caa aca ctc tag 1086
 Thr Phe Thr Pro Ser Ser Gln Thr Leu *
 355 360

<210> 10
 <211> 361
 <212> PRT
 <213> E. coli

<400> 10

Met Asn Lys Ile Leu Phe Ile Phe Thr Leu Phe Phe Ser Ser Val Leu
 1 5 10 15
 Phe Thr Phe Ala Val Ser Ala Asp Lys Ile Pro Gly Asp Glu Ser Ile
 20 25 30
 Thr Asn Ile Phe Gly Pro Arg Asp Arg Asn Glu Ser Ser Pro Lys His
 35 40 45
 Asn Ile Leu Asn Asn His Ile Thr Ala Tyr Ser Glu Ser His Thr Leu
 50 55 60
 Tyr Asp Arg Met Thr Phe Leu Cys Leu Ser Ser His Asn Thr Leu Asn
 65 70 75 80
 Gly Ala Cys Pro Thr Ser Glu Asn Pro Ser Ser Ser Ser Val Ser Gly
 85 90 95
 Glu Thr Asn Ile Thr Leu Gln Phe Thr Glu Lys Arg Ser Leu Ile Lys
 100 105 110
 Arg Glu Leu Gln Ile Lys Gly Tyr Lys Gln Leu Leu Phe Lys Ser Val
 115 120 125
 Asn Cys Pro Ser Gly Leu Thr Leu Asn Ser Ala His Phe Asn Cys Asn
 130 135 140
 Lys Asn Ala Ala Ser Gly Ala Ser Leu Tyr Leu Tyr Ile Pro Ala Gly
 145 150 155 160
 Glu Leu Lys Asn Leu Pro Phe Gly Gly Ile Trp Asp Ala Thr Leu Lys
 165 170 175
 Leu Arg Val Lys Arg Arg Tyr Ser Glu Thr Tyr Gly Thr Tyr Ile
 180 185 190
 Asn Ile Thr Ile Lys Leu Thr Asp Lys Gly Asn Ile Gln Ile Trp Leu
 195 200 205
 Pro Gln Phe Lys Ser Asp Ala Arg Val Asp Leu Asn Leu Arg Pro Thr
 210 215 220
 Gly Gly Gly Thr Tyr Ile Gly Arg Asn Ser Val Asp Met Cys Phe Tyr
 225 230 235 240
 Asp Gly Tyr Ser Thr Asn Ser Ser Ser Leu Glu Ile Arg Phe Gln Asp
 245 250 255
 Asn Asn Pro Lys Ser Asp Gly Lys Phe Tyr Leu Arg Lys Ile Asn Asp
 260 265 270

Asp Thr Lys Glu Ile Ala Tyr Thr Leu Ser Leu Leu Leu Ala Gly Lys
 275 280 285
 Ser Leu Thr Pro Thr Asn Gly Thr Ser Leu Asn Ile Ala Asp Ala Ala
 290 295 300
 Ser Leu Glu Thr Asn Trp Asn Arg Ile Thr Ala Val Thr Met Pro Glu
 305 310 315 320
 Ile Ser Val Pro Val Leu Cys Trp Pro Gly Arg Leu Gln Leu Asp Ala
 325 330 335
 Lys Val Glu Asn Pro Glu Ala Gly Gln Tyr Met Gly Asn Ile Asn Val
 340 345 350
 Thr Phe Thr Pro Ser Ser Gln Thr Leu
 355 360

<210> 11
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR Primer

<400> 11
 gttgacccta caattgatat tttgcaagc

29

<210> 12
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR Primer

<400> 12
 cgacccact ataattcccg ccgttggtgc

30

<210> 13
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR Primer

<400> 13
 gtgatatgtt ttgttcactt ggtaaagatc

30

<210> 14
 <211> 36
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR Primer

<400> 14
 ctcatggctc catttggtgc aaatgcaaac tttatg 36

 <210> 15
 <211> 60
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR Primer

 <400> 15
 gggatcgatc ccggggcggc cgcgggcccgtgtaccaggcc ttctagaaag cttgacgtcg 60

 <210> 16
 <211> 61
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR Primer

 <400> 16
 cccgctagcg gcgcgcctcg cgaggatccg tcgacgacgt caagctttct agaaggcctg 60
 g 61

 <210> 17
 <211> 20
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR Primer

 <400> 17
 aagcttgacg tcgtcgacgg 20

 <210> 18
 <211> 22
 <212> DNA
 <213> Artificial Sequence

 <220>
 <223> PCR Primer

 <400> 18
 cccgctagcg gcgcgcctcg cg 22

 <210> 19
 <211> 26
 <212> DNA
 <213> Artificial Sequence

 <220>

<223> PCR Primer

<400> 19

ccgtgctgac tctacacccc cagatg

26

<210> 20

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 20

gcacatagag aggatagtaa cgccg

25

<210> 21

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 21

cggtcattgt tggccgtgcg ctgcc

25

<210> 22

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 22

cacgcagcgc gctgatgcct tccacgcg

28

<210> 23

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 23

catatttgat atctgagata tctgg

25

<210> 24

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 24

tgttgcattc agattgaacg gag

23

<210> 25

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 25

tattatgatt cataaataca ctgt

24

<210> 26

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR Primer

<400> 26

tgtgggtatt tgtttggaca tcgcagcatt aaatataaaa atagcacagg

50

<210> 27

<211> 7239

<212> DNA

<213> E. coli

<220>

<221> CDS

<222> (283) ... (999)

<221> CDS

<222> (1028) ... (1531)

<221> CDS

<222> (1589) ... (4192)

<221> CDS

<222> (4196) ... (5281)

<221> CDS

<222> (5790) ... (6119)

<400> 27

atatatctta ttgaggaata tcggtgtcat tgagtaccgt taacttaaga taaagaatct 60
gtctggaaat cgcaggacca agaactctca gtacatctgt ggcgataata ttatcgcttc 120
ttatacattc caatatgcag ttcttgtggg tatttgtttg gacatcgag cattaaatat 180
aaaaatagca caggaggcat aattatttgt ttttactgtc ttattttttt atcccathtt 240
tttttgtttt gatttatctt tgatgaaagc tcaggaggga atatg cat aaa tta ttt 297

His Lys Leu Phe

tgt Cys 5	tta Leu	cta Leu	agt Ser	tta Leu	ctc Leu	ata Ile	act Thr	cca Pro	ttt Phe	gtt Val	gca Ala	aat Asn	gca Ala	aac Asn	ttt Phe	345
101520																
atg Met	ata Ile	tat Tyr	cca Pro	ata Ile	tca Ser	aaa Lys	gat Asp	tta Leu	aag Lys	aat Asn	gga Gly	aat Asn	agc Ser	gag Glu	tta Leu	393
253035																
att Ile	cgt Arg	gtt Val	tat Tyr	tca Ser	aaa Lys	tca Ser	aaa Lys	gag Glu	ata Ile	caa Gln	tat Tyr	ata Ile	aaa Lys	ata Ile	tat Tyr	441
404550																
aca Thr	aaa Lys	aag Lys	att Ile	att Ile	aat Asn	ccc Pro	ggc Gly	aca Thr	act Thr	gaa Glu	gaa Glu	cat His	gaa Glu	gtt Val	gat Asp	489
556065																
atg Met	ccc Pro	aat Asn	tgg Trp	gat Asp	ggg Gly	ggg Gly	ttt Phe	gta Val	gtt Val	act Thr	cct Pro	caa Gln	aaa Lys	gtt Val	att Ile	537
707580																
ctt Leu	cct Pro	gca Ala	gga Gly	ggg Gly	agt Ser	aaa Lys	tca Ser	ata Ile	cgt Arg	tta Leu	act Thr	caa Gln	ttt Phe	aga Arg	ata Ile	585
859095100																
cca Pro	aaa Lys	aaa Lys	gag Glu	gaa Glu	att Ile	tat Tyr	aga Arg	gta Val	tat Tyr	ttt Phe	gag Glu	gcg Ala	gta Val	aaa Lys	cca Pro	633
105110115																
gat Asp	agc Ser	aaa Lys	gaa Glu	aat Asn	gta Val	att Ile	gat Asp	aat Asn	aaa Lys	aaa Lys	cta Leu	aca Thr	aca Thr	gag Glu	cta Leu	681
120125130																
tct Ser	gtt Val	aat Asn	ata Ile	att Ile	tat Tyr	gcg Ala	gct Ala	cta Leu	atc Ile	aga Arg	tct Ser	tta Leu	cca Pro	agt Ser	gaa Glu	729
135140145																
caa Gln	aac Asn	ata Ile	tca Ser	cta Leu	aac Asn	att Ile	tct Ser	aga Arg	aat Asn	gca Ala	aga Arg	aaa Lys	aat Asn	ata Ile	att Ile	777
150155160																
att Ile	tat Tyr	aat Asn	aat Asn	ggg Gly	aat Asn	gtt Val	aga Arg	gca Ala	ggg Gly	gtt Val	aaa Lys	gat Asp	att Ile	tat Tyr	ttt Phe	825
165170175180																
tgt Cys	aag Lys	tca Ser	tct Ser	aat Asn	atc Ile	gat Asp	gat Asp	agc Ser	tgt Cys	gta Val	aaa Lys	aaa Lys	acg Thr	cat His	aac Asn	873
185190195																
aag Lys	aat Asn	ata Ile	tat Tyr	cca Pro	gaa Glu	aag Lys	tca Ser	ttt Phe	gat Asp	acg Thr	ctg Leu	gtt Val	aat Asn	aac Asn	aat Asn	921
200205210																
ttt	tct	tat	qtt	ttc	att	aaa	tta	aac	cat	gaa	gac	ata	gaa	aaa	gag	969

Phe Ser Tyr Val Phe Ile Lys Leu Asn His Glu Asp Ile Glu Lys Glu
215 220 225

caa gga cta ata caa tta aaa gtt cct tga tta ctcatctata tactaaggag 1022
Gln Gly Leu Ile Gln Leu Lys Val Pro * Leu
230 235

ttctaatagaa attaaaaaaaa actattggtg caatg gca ctg acc aca atg ttt 1075
Ala Leu Thr Thr Met Phe
240

gta gct atg agt gct tct gca gta gag aaa aat atc act gta aca gct 1123
Val Ala Met Ser Ala Ser Ala Val Glu Lys Asn Ile Thr Val Thr Ala
245 250 255 260

agt gtt gat cct aca att gat att ttg caa gct gat ggt agt agt tta 1171
Ser Val Asp Pro Thr Ile Asp Ile Leu Gln Ala Asp Gly Ser Ser Leu
265 270 275

cct act gct gta gaa tta acc tat tca cct gcg gca agt cgt ttt gaa 1219
Pro Thr Ala Val Glu Leu Thr Tyr Ser Pro Ala Ala Ser Arg Phe Glu
280 285 290

aat tat aaa atc gca act aaa gtt cat aca aat gtt ata aat aaa aat 1267
Asn Tyr Lys Ile Ala Thr Lys Val His Thr Asn Val Ile Asn Lys Asn
295 300 305

gta cta gtt aag ctt gta aat gat cca aaa ctt aca aat gtt ttg gat 1315
Val Leu Val Lys Leu Val Asn Asp Pro Lys Leu Thr Asn Val Leu Asp
310 315 320

tct aca aaa caa ctc ccc att act gta tca tat gga gga aag act cta 1363
Ser Thr Lys Gln Leu Pro Ile Thr Val Ser Tyr Gly Gly Lys Thr Leu
325 330 335 340

tca acc gca gat gtg act ttt gaa cct gca gaa tta aat ttt gga acg 1411
Ser Thr Ala Asp Val Thr Phe Glu Pro Ala Glu Leu Asn Phe Gly Thr
345 350 355

tca ggt gta act ggt gta tct tct tcc caa gat tta gtg att ggt gcg 1459
Ser Gly Val Thr Gly Val Ser Ser Ser Gln Asp Leu Val Ile Gly Ala
360 365 370

act aca gca caa gca cca acg gcg gga aat tat agt ggg gtc gtt tct 1507
Thr Thr Ala Gln Ala Pro Thr Ala Gly Asn Tyr Ser Gly Val Val Ser
375 380 385

atc tta atg acc tta gca tca taa ata ttttaatat taaaggagca 1554
Ile Leu Met Thr Leu Ala Ser * Ile
390 395

ggcacactgc tccttattat atggcaataa taaaatg aca aaa aaa aat aca tta 1609
Thr Lys Lys Asn Thr Leu
400

tat	ata	acg	atc	atc	gca	atg	cta	act	cca	tat	tca	ggt	ttt	tcc	gga	1657
Tyr	Ile	Thr	Ile	Ile	Ala	Met	Leu	Thr	Pro	Tyr	Ser	Val	Phe	Ser	Gly	
		405					410					415				
gat	ata	ccc	aac	tct	ttc	cgt	gat	tta	tgg	gga	gaa	caa	gat	gaa	ttt	1705
Asp	Ile	Pro	Asn	Ser	Phe	Arg	Asp	Leu	Trp	Gly	Glu	Gln	Asp	Glu	Phe	
		420				425					430					
tat	gaa	gta	aaa	cta	tat	gga	caa	act	cta	gga	ata	cat	cga	att	aaa	1753
Tyr	Glu	Val	Lys	Leu	Tyr	Gly	Gln	Thr	Leu	Gly	Ile	His	Arg	Ile	Lys	
					440					445					450	
aca	acc	cca	aca	cat	att	aag	ttt	tat	tca	ccc	gaa	agc	att	tta	gat	1801
Thr	Thr	Pro	Thr	His	Ile	Lys	Phe	Tyr	Ser	Pro	Glu	Ser	Ile	Leu	Asp	
				455					460					465		
aaa	ata	aat	gta	aaa	aaa	gaa	aag	gaa	aag	aaa	ttg	agt	ggt	ttg	ttc	1849
Lys	Ile	Asn	Val	Lys	Lys	Glu	Lys	Glu	Lys	Lys	Leu	Ser	Val	Leu	Phe	
			470					475					480			
act	aat	tct	ttt	tca	aga	aat	ggc	aat	atg	agt	tgt	cag	ggg	aat	gct	1897
Thr	Asn	Ser	Phe	Ser	Arg	Asn	Gly	Asn	Met	Ser	Cys	Gln	Gly	Asn	Ala	
		485					490					495				
act	ata	cag	tat	aac	tgc	aat	tac	att	aaa	aca	aaa	tca	gta	gat	gtc	1945
Thr	Ile	Gln	Tyr	Asn	Cys	Asn	Tyr	Ile	Lys	Thr	Lys	Ser	Val	Asp	Val	
		500				505					510					
atc	ggt	gat	gat	ggt	gat	aat	ggt	ggt	aac	ctt	ttt	ata	ggt	aat	gaa	1993
Ile	Val	Asp	Asp	Val	Asp	Asn	Val	Val	Asn	Leu	Phe	Ile	Gly	Asn	Glu	
					520					525					530	
ttt	ctg	gat	tct	gaa	gca	cac	aat	gat	gaa	tat	cat	caa	tta	tca	cga	2041
Phe	Leu	Asp	Ser	Glu	Ala	His	Asn	Asp	Glu	Tyr	His	Gln	Leu	Ser	Arg	
				535					540					545		
aat	gta	aaa	aaa	gct	ttt	ata	caa	agc	cag	aca	att	aat	gtc	tca	gat	2089
Asn	Val	Lys	Lys	Ala	Phe	Ile	Gln	Ser	Gln	Thr	Ile	Asn	Val	Ser	Asp	
			550					555					560			
tct	ggg	aag	tat	aaa	agt	ttg	tct	ggt	tca	ggg	aat	agc	gcg	ctg	ggt	2137
Ser	Gly	Lys	Tyr	Lys	Ser	Leu	Ser	Val	Ser	Gly	Asn	Ser	Ala	Leu	Gly	
		565					570					575				
att	aca	gat	aca	agt	tat	gct	gtc	tta	aat	tgg	tgg	atg	aat	tac	aat	2185
Ile	Thr	Asp	Thr	Ser	Tyr	Ala	Val	Leu	Asn	Trp	Trp	Met	Asn	Tyr	Asn	
		580				585					590					
aaa	ttt	aat	ggt	tac	agc	aac	aac	gaa	aga	aca	atc	aat	agt	ttg	tac	2233
Lys	Phe	Asn	Gly	Tyr	Ser	Asn	Asn	Glu	Arg	Thr	Ile	Asn	Ser	Leu	Tyr	
					600				605						610	
ttt	aga	cat	gat	tta	gat	aag	aga	tat	tat	tat	caa	ttt	gga	cga	atg	22

gat cgt aca gat ttg tca caa agt att agc ggg aac ttt aat ttt aac	2329
Asp Arg Thr Asp Leu Ser Gln Ser Ile Ser Gly Asn Phe Asn Phe Asn	
630 635 640	
tta ctt cct tta ccc gat att gat ggt ata agg aca gga acc aca caa	2377
Leu Leu Pro Leu Pro Asp Ile Asp Gly Ile Arg Thr Gly Thr Thr Gln	
645 650 655	
tct tat atc aaa aat aca gat aag ttt atc gca tcc cct gta act gtt	2425
Ser Tyr Ile Lys Asn Thr Asp Lys Phe Ile Ala Ser Pro Val Thr Val	
660 665 670	
atg tta act aat ttt tcc aga gtg gaa gct ttt cgc aat aat caa tta	2473
Met Leu Thr Asn Phe Ser Arg Val Glu Ala Phe Arg Asn Asn Gln Leu	
675 680 685 690	
ttg ggc gta tgg tat tta gat tct gga gta aat gaa tta gat aca gct	2521
Leu Gly Val Trp Tyr Leu Asp Ser Gly Val Asn Glu Leu Asp Thr Ala	
695 700 705	
cgt tta cct tat ggt agt tac gat ctt aaa tta aaa att ttt gaa aat	2569
Arg Leu Pro Tyr Gly Ser Tyr Asp Leu Lys Leu Lys Ile Phe Glu Asn	
710 715 720	
act cag tta gtt cgt gaa gaa ata att cct ttt aat aaa ggg aga agt	2617
Thr Gln Leu Val Arg Glu Glu Ile Ile Pro Phe Asn Lys Gly Arg Ser	
725 730 735	
tct att ggt gat atg caa tgg gac gtt ttc att cag gga ggg aat att	2665
Ser Ile Gly Asp Met Gln Trp Asp Val Phe Ile Gln Gly Gly Asn Ile	
740 745 750	
att aat gac aag gat cgt tac ata gaa aaa caa aat aat cat aag tca	2713
Ile Asn Asp Lys Asp Arg Tyr Ile Glu Lys Gln Asn Asn His Lys Ser	
755 760 765 770	
tca gtt aat gct ggg cta cgt tta cca att acg aaa aat atc tct gtt	2761
Ser Val Asn Ala Gly Leu Arg Leu Pro Ile Thr Lys Asn Ile Ser Val	
775 780 785	
caa caa gga gca tct gtt ata gat aat aaa aat tat tat gaa ggg agt	2809
Gln Gln Gly Ala Ser Val Ile Asp Asn Lys Asn Tyr Tyr Glu Gly Ser	
790 795 800	
ctc aaa tgg aat tcc ggc att ctg tct ggc tca cta aat agt gag ttc	2857
Leu Lys Trp Asn Ser Gly Ile Leu Ser Gly Ser Leu Asn Ser Glu Phe	
805 810 815	
agt ttt ctt tgg gga gat aat gca aaa ggt aat tat caa agt atc tcg	2905
Ser Phe Leu Trp Gly Asp Asn Ala Lys Gly Asn Tyr Gln Ser Ile Ser	
820 825 830	
tat acc gat gga ttt agt tta tca ttt tat cat aat gat aag cgg gtc	2953
Tyr Thr Asp Gly Phe Ser Leu Ser Phe Tyr His Asn Asp Lys Arg Val	

835	840	845	850	
gat aat tgt gga aga aat tac aat gct ggt tgg agt gga tgc tac gaa	Asp Asn Cys Gly Arg Asn Tyr Asn Ala Gly Trp Ser Gly Cys Tyr Glu	3001		
	855 860 865			
tca tat tcg gca tct tta agt att cct tta ttg gga tgg aca agt act	Ser Tyr Ser Ala Ser Leu Ser Ile Pro Leu Leu Gly Trp Thr Ser Thr	3049		
	870 875 880			
ctg gga tat agt gac act tat agt gaa tca gtt tat aaa aac cat att	Leu Gly Tyr Ser Asp Thr Tyr Ser Glu Ser Val Tyr Lys Asn His Ile	3097		
	885 890 895			
ctt tct gaa tat ggt ttt tat aat caa aac ata tat aaa ggg aga acc	Leu Ser Glu Tyr Gly Phe Tyr Asn Gln Asn Ile Tyr Lys Gly Arg Thr	3145		
	900 905 910			
caa aga tgg caa ctg act tcg tcc acc tct tta aaa tgg atg gat tat	Gln Arg Trp Gln Leu Thr Ser Ser Thr Ser Leu Lys Trp Met Asp Tyr	3193		
	915 920 925 930			
aat ttt atg cca gca att gga ata tat aac agt gag caa aga caa ctg	Asn Phe Met Pro Ala Ile Gly Ile Tyr Asn Ser Glu Gln Arg Gln Leu	3241		
	935 940 945			
act gat aaa ggc gga tat ata tct gta act ctc acc cga gcc agc aga	Thr Asp Lys Gly Gly Tyr Ile Ser Val Thr Leu Thr Arg Ala Ser Arg	3289		
	950 955 960			
gaa aat tca tta aac gca ggg tat tct tac aac tat tcc aga gga aag	Glu Asn Ser Leu Asn Ala Gly Tyr Ser Tyr Asn Tyr Ser Arg Gly Lys	3337		
	965 970 975			
tat tct tct aac gaa tta ttt gtt gat gga tat atg aca tca aca aat	Tyr Ser Ser Asn Glu Leu Phe Val Asp Gly Tyr Met Thr Ser Thr Asn	3385		
	980 985 990			
aat ggt gac tat cat gag gta aga atg cgt ttt aat aaa aat aga cat	Asn Gly Asp Tyr His Glu Val Arg Met Arg Phe Asn Lys Asn Arg His	3433		
	995 1000 1005 1010			
aat gca gaa ggt aga ctt tca ggt cgt ata aac aat cga ttt gga gat	Asn Ala Glu Gly Arg Leu Ser Gly Arg Ile Asn Asn Arg Phe Gly Asp	3481		
	1015 1020 1025			
tta aat ggt tca ttc agc atg aat aaa aac aga aac acc aac agt agc	Leu Asn Gly Ser Phe Ser Met Asn Lys Asn Arg Asn Thr Asn Ser Ser	3529		
	1030 1035 1040			
aat cat tct ctc act ggt ggt tat aat tcc tca ttt gct ctt aca agt	Asn His Ser Leu Thr Gly Gly Tyr Asn Ser Ser Phe Ala Leu Thr Ser	3577		
	1045 1050 1055			
gat gga ttt tac tgg gga gga agt gca tct ggt ttg aca aaa cta gct		3625		

Asp	Gly	Phe	Tyr	Trp	Gly	Ser	Ala	Ser	Gly	Leu	Thr	Lys	Leu	Ala			
1060					1065					1070							
ggc	ggt	att	atc	aag	gtt	aaa	tca	aac	gat	act	aaa	aaa	aat	ctg	gta	3673	
Gly	Gly	Ile	Ile	Lys	Val	Lys	Ser	Asn	Asp	Thr	Lys	Lys	Asn	Leu	Val		
1075					1080					1085					1090		
aaa	gtg	act	ggg	gca	ttg	tac	ggt	gat	tat	tcg	cta	ggg	agc	aac	gat	3721	
Lys	Val	Thr	Gly	Ala	Leu	Tyr	Gly	Asp	Tyr	Ser	Leu	Gly	Ser	Asn	Asp		
				1095					1100					1105			
aat	gct	ttt	att	cct	gta	cca	gca	tta	act	cca	gcc	agt	tta	att	att	3769	
Asn	Ala	Phe	Ile	Pro	Val	Pro	Ala	Leu	Thr	Pro	Ala	Ser	Leu	Ile	Ile		
				1110					1115					1120			
gaa	gat	aat	aat	tat	ggt	gac	aag	aat	att	tct	gta	ctt	gca	cca	acg	3817	
Glu	Asp	Asn	Asn	Tyr	Gly	Asp	Lys	Asn	Ile	Ser	Val	Leu	Ala	Pro	Thr		
				1125					1130					1135			
aac	aac	gat	atg	ttt	ata	ttg	ccg	ggt	aat	ggt	tat	cct	ggt	gaa	att	3865	
Asn	Asn	Asp	Met	Phe	Ile	Leu	Pro	Gly	Asn	Val	Tyr	Pro	Val	Glu	Ile		
				1140					1145					1150			
gaa	acc	aaa	gta	agt	gtt	tct	tat	att	ggt	aga	ggt	ttt	gac	aaa	aac	3913	
Glu	Thr	Lys	Val	Ser	Val	Ser	Tyr	Ile	Gly	Arg	Gly	Phe	Asp	Lys	Asn		
1155					1160					1165					1170		
ggc	acg	cca	ctt	tct	ggc	gca	cat	ggt	ttg	aat	gaa	cca	cat	ggt	atc	3961	
Gly	Thr	Pro	Leu	Ser	Gly	Ala	His	Val	Leu	Asn	Glu	Pro	His	Val	Ile		
				1175					1180					1185			
ctg	gat	gag	gac	ggt	gga	ttt	tcg	ttt	gaa	tat	aca	ggt	aat	gag	aaa	4009	
Leu	Asp	Glu	Asp	Gly	Gly	Phe	Ser	Phe	Glu	Tyr	Thr	Gly	Asn	Glu	Lys		
				1190					1195					1200			
aca	ctt	ttt	tta	tta	aag	ggc	aga	act	att	tat	aca	tgt	caa	ctg	ggg	4057	
Thr	Leu	Phe	Leu	Leu	Lys	Gly	Arg	Thr	Ile	Tyr	Thr	Cys	Gln	Leu	Gly		
1205					1210					1215							
aaa	aat	aaa	gtt	cac	aaa	ggc	att	ggt	ttc	gtc	gga	gat	ggt	ata	tgt	4105	
Lys	Asn	Lys	Val	His	Lys	Gly	Ile	Val	Phe	Val	Gly	Asp	Val	Ile	Cys		
1220					1225					1230							
gat	gtt	aat	agc	aca	agt	tcc	tta	cca	gat	gaa	ttt	gta	aag	aac	cca	4153	
Asp	Val	Asn	Ser	Thr	Ser	Ser	Leu	Pro	Asp	Glu	Phe	Val	Lys	Asn	Pro		
1235					1240					1245					1250		
cgt	gtg	cag	gat	ttg	ctg	gca	aag	aat	gat	aaa	gga	taa	acg			4195	
Arg	Val	Gln	Asp	Leu	Leu	Ala	Lys	Asn	Asp	Lys	Gly	*	Thr				
				1255					1260								
atg	aat	aag	att	tta	ttt	att	ttt	aca	ttg	ttt	ttc	tct	tca	gta	ctt	4243	
	Asn	Lys	Ile	Leu	Phe	Ile	Phe	Thr	Leu	Phe	Phe	Ser	Ser	Val	Leu		
1265					1270					1275							

ttt aca ttt gct gta tcg gca gat aaa att ccc gga gat gaa agc ata	4291
Phe Thr Phe Ala Val Ser Ala Asp Lys Ile Pro Gly Asp Glu Ser Ile	
1280 1285 1290	
act aat att ttt ggc ccg cgt gac agg aac gaa tct tcc ccc aaa cat	4339
Thr Asn Ile Phe Gly Pro Arg Asp Arg Asn Glu Ser Ser Pro Lys His	
1295 1300 1305 1310	
aat ata tta aat aac cat att aca gca tac agt gaa agt cat act ctg	4387
Asn Ile Leu Asn Asn His Ile Thr Ala Tyr Ser Glu Ser His Thr Leu	
1315 1320 1325	
tat gat agg atg act ttt tta tgt ttg tct tct cac aat aca ctt aat	4435
Tyr Asp Arg Met Thr Phe Leu Cys Leu Ser Ser His Asn Thr Leu Asn	
1330 1335 1340	
gga gca tgt cca acc agt gag aat cct agc agt tca tcg gtc agc ggt	4483
Gly Ala Cys Pro Thr Ser Glu Asn Pro Ser Ser Ser Ser Val Ser Gly	
1345 1350 1355	
gaa aca aat ata aca tta caa ttt acg gaa aaa aga agt tta ata aaa	4531
Glu Thr Asn Ile Thr Leu Gln Phe Thr Glu Lys Arg Ser Leu Ile Lys	
1360 1365 1370	
aga gag cta caa att aaa ggc tat aaa caa tta ttg ttc aaa agt gtt	4579
Arg Glu Leu Gln Ile Lys Gly Tyr Lys Gln Leu Leu Phe Lys Ser Val	
1375 1380 1385 1390	
aac tgc cca tcc ggc cta aca ctt aac tca gct cat ttt aac tgt aat	4627
Asn Cys Pro Ser Gly Leu Thr Leu Asn Ser Ala His Phe Asn Cys Asn	
1395 1400 1405	
aaa aac gcg gct tca ggt gca agt tta tat tta tat att cct gct ggc	4675
Lys Asn Ala Ala Ser Gly Ala Ser Leu Tyr Leu Tyr Ile Pro Ala Gly	
1410 1415 1420	
gaa cta aaa aat ttg cct ttt ggt ggt atc tgg gat gct act ctg aag	4723
Glu Leu Lys Asn Leu Pro Phe Gly Gly Ile Trp Asp Ala Thr Leu Lys	
1425 1430 1435	
tta aga gta aaa aga cga tat agt gag acc tat gga act tac act ata	4771
Leu Arg Val Lys Arg Arg Tyr Ser Glu Thr Tyr Gly Thr Tyr Thr Ile	
1440 1445 1450	
aat atc act att aaa tta act gat aag gga aat att cag ata tgg tta	4819
Asn Ile Thr Ile Lys Leu Thr Asp Lys Gly Asn Ile Gln Ile Trp Leu	
1455 1460 1465 1470	
cct cag ttc aaa agt gac gct cgc gtc gat ctt aac ttg cgt cca act	4867
Pro Gln Phe Lys Ser Asp Ala Arg Val Asp Leu Asn Leu Arg Pro Thr	
1475 1480 1485	
ggt ggg ggc aca tat att gga aga aat tct gtt gat atg tgc ttt tat	4915
Gly Gly Gly Thr Tyr Ile Gly Arg Asn Ser Val Asp Met Cys Phe Tyr	
1490 1495 1500	

gat gga tat agt act aac agc agc tct ttg gag ata aga ttt cag gat 4963
Asp Gly Tyr Ser Thr Asn Ser Ser Ser Leu Glu Ile Arg Phe Gln Asp
1505 1510 1515

aac aat cct aaa tct gat ggg aaa ttt tat cta agg aaa ata aat gat 5011
Asn Asn Pro Lys Ser Asp Gly Lys Phe Tyr Leu Arg Lys Ile Asn Asp
1520 1525 1530

gac acc aaa gaa att gca tat act ttg tca ctt ctc ttg gcg ggt aaa 5059
Asp Thr Lys Glu Ile Ala Tyr Thr Leu Ser Leu Leu Leu Ala Gly Lys
1535 1540 1545 1550

agt tta act cca aca aat gga acg tca tta aat att gct gac gca gct 5107
Ser Leu Thr Pro Thr Asn Gly Thr Ser Leu Asn Ile Ala Asp Ala Ala
1555 1560 1565

tct ctg gaa aca aac tgg aat aga att aca gct gtc acc atg cca gaa 5155
Ser Leu Glu Thr Asn Trp Asn Arg Ile Thr Ala Val Thr Met Pro Glu
1570 1575 1580

atc agt gtt ccg gtg ttg tgt tgg cct gga cgt ttg caa ttg gat gca 5203
Ile Ser Val Pro Val Leu Cys Trp Pro Gly Arg Leu Gln Leu Asp Ala
1585 1590 1595

aaa gtg gaa aat ccc gag gct gga caa tat atg ggt aat att aat gtt 5251
Lys Val Glu Asn Pro Glu Ala Gly Gln Tyr Met Gly Asn Ile Asn Val
1600 1605 1610

act ttc aca cca agt agt caa aca ctc tag ata acaacaatat tggcgctatt 5304
Thr Phe Thr Pro Ser Ser Gln Thr Leu * Ile
1615 1620

gcgcgcgaat attgtaaagg ggtaattctgt ttgttaacaa aacattttgt ttcaattcag 5364
tttgcacgc aataaatctc tactagagac atttttatac agcatagtat tatacaacac 5424
attcaaaaata aggatatttt tatccacctt taaaataagt aaaaaactgc tttggtataa 5484
caccataatg tttattaaaa accctaataa aataagatgt actggaaatt ccaatcatat 5544
ttgatattctg agatatctgg tatgaatttt caagtagtaa taacgctgcc ttgctcattc 5604
tcaattgcat taagaactgg ttaaaattag tattctcaga ttctagtctt tttctgatgg 5664
ttatttctga ttcattaaac atatctgcaa tgatagccag tgtccatttt ctggatagat 5724
ctttttcgat aatatttctg accttgctcag aaaaaaattc acagatgata tataaattga 5784
ttctatta ttt ttt att att ctc gat ctt tga aat taa ata tat caa att 5834
Phe Phe Ile Ile Leu Asp Leu * Asn * Ile Tyr Gln Ile
1625 1630 1635

aga tat aaa agc tga gtc atc ata gct att tat att ttt taa tac atc 5882
Arg Tyr Lys Ser * Val Ile Ile Ala Ile Tyr Ile Phe * Tyr Ile
1640 1645 1650

cag taa ggt ttt atc cac ttc tgt ttt cat tat ttt cct tga cat att 5930
Gln * Gly Phe Ile His Phe Cys Phe His Tyr Phe Pro * His Ile
1655 1660

tct aca atc att ggt atc tat ttt tga cat acc ata tat tat cat caa 5978
Ser Thr Ile Ile Gly Ile Tyr Phe * His Thr Ile Tyr Tyr His Gln

1665

1670

1675

tgc atc ctt taa atg tct tag tat gtc tcc gtt caa tct gaa tgc aac 6026
 Cys Ile Leu * Met Ser * Tyr Val Ser Val Gln Ser Glu Cys Asn
 1680 1685 1690

ata tgg ttt ttc tga taa aat ttg ctt ctg tat tct tac aga tat att 6074
 Ile Trp Phe Phe * * Asn Leu Leu Leu Tyr Ser Tyr Arg Tyr Ile
 1695 1700 1705

cac ccc tct ttc aag aaa tac agg t gatgctgccca acttactgat 6119
 His Pro Ser Phe Lys Lys Tyr Arg
 1710 1715

ttagtgatg atggtgtttt tgaggtgctc cagtggcttc tgtttctatc agctgtccct 6179
 cctgttcagc tactgacggg gtggtgctga acggcaaaag cactgccgga catcagcgct 6239
 atctctgctc tctactgccgt aaaacatggc aactgcagtt cacttacact gcttctcaac 6299
 ccggtacgca ccagaaaatc attgatattg ccatgaatgg cgttggatgc cgggcaacag 6359
 cccgcattat gggcggtggc ctcaacacga ttttacgtca cttaaaaaac tcaggccgca 6419
 gtcggttaacc tcgcgcatac agccgggcag tgacgtcatc gtctgcgcgg aaatggacga 6479
 acagtggggc tatgtcgggg ctaaatcgcg ccagcgctgg ctgttttacg cgatgacag 6539
 gctccggaag acggttggtg cgcacgtatt cgggtaacgc actatggcga cgctggggcg 6599
 tcttatgagc ctgctgtcac cctttgacgt ggtgatattg atgacggatg gctggccgct 6659
 gtatgaatcc cgcctgaagg gaaagctgca cgtaatcagc aagcgatata cgcagcgaat 6719
 tgagcggcat aacctgaatc tgaggcagca cctggcacgg ctgggacgga agtcgctgtc 6779
 gttctcaaaa tcggtggagc tgcatgacaa agtcacggg cattatctga acataaaaca 6839
 ctactcccgc gttagagtca ttacctggtt cactgtattat tatccgtgac tctttcctgg 6899
 taactcccgc ataataacct cacttttcca gtattccaga agatgatgtt ttttcctcga 6959
 taataaaaaat gtgccaatat ggaaataaga aatcggattt tttatcagca tacgcaaatt 7019
 ttcagataac aatgaataca gatgtatttt atatacacag ataaaaccgc gcaacagaca 7079
 taaatatgac agtagcatga aaaagcagag agagacaggg tgatacagaa aagtaactat 7139
 ttttttagct atagtattat tggttttacc tattttcgtg attgtgtttc tgtatatttg 7199
 acaatgagtc tctcagaatc ggtttctcga agtgacgagc 7239

<210> 28

<211> 361

<212> PRT

<213> Artificial Sequence

<220>

<223> ETEC Protein Homology Sequence

<400> 28

Thr Met Asn Lys Ile Leu Phe Ile Phe Thr Leu Phe Phe Ser Ser Val
 1 5 10 15
 Leu Phe Thr Phe Ala Val Ser Ala Asp Lys Ile Pro Gly Asp Glu Ser
 20 25 30
 Ile Thr Asn Ile Phe Gly Pro Arg Asp Arg Asn Glu Ser Ser Pro Lys
 35 40 45
 His Asn Ile Leu Asn Asn His Ile Thr Ala Tyr Ser Glu Ser His Thr
 50 55 60
 Leu Tyr Asp Arg Met Thr Phe Leu Cys Leu Ser Ser His Asn Thr Leu
 65 70 75 80
 Asn Gly Ala Cys Pro Thr Ser Glu Asn Pro Ser Ser Ser Ser Val Ser
 85 90 95

SECRET

<211> 359

<213> Artificial Sequence

<223> ETEC Protein Homology Sequence

Met	Asn	Lys	Ile	Leu	Phe	Ile	Phe	Thr	Leu	Phe	Phe	Ser	Ser	Gly	Phe
1				5					10					15	
Phe	Thr	Phe	Ala	Val	Ser	Ala	Asp	Lys	Asn	Pro	Gly	Ser	Glu	Asn	Met
			20					25					30		
Thr	Asn	Thr	Ile	Gly	Pro	His	Asp	Arg	Gly	Gly	Ser	Ser	Pro	Ile	Tyr
		35					40					45			
Asn	Ile	Leu	Asn	Ser	Tyr	Leu	Thr	Ala	Tyr	Asn	Gly	Ser	His	His	Leu
	50					55					60				
Tyr	Asp	Arg	Met	Ser	Phe	Leu	Cys	Leu	Ser	Ser	Gln	Asn	Thr	Leu	Asn
65					70					75					80

Gly	Ala	Cys	Pro	Ser	Asp	Ala	Pro	Gly	Thr	Ala	Thr	Ile	Asp	Gly		
				85					90					95	Gly	
Glu	Thr	Asn	Ile	Thr	Leu	Gln	Phe	Thr	Glu	Lys	Arg	Ser	Leu	Ile	Lys	
				100					105					110	Lys	
Arg	Glu	Leu	Gln	Ile	Lys	Gly	Tyr	Lys	Gln	Phe	Leu	Phe	Lys	Asn	Ala	
				115					120					125	Ala	
Asn	Cys	Pro	Ser	Lys	Leu	Ala	Leu	Asn	Ser	Ser	His	Phe	Gln	Cys	Asn	
				130					135					140	Asn	
Arg	Glu	Gln	Ala	Ser	Gly	Ala	Thr	Leu	Ser	Leu	Tyr	Ile	Pro	Ala	Gly	
145					150					155					160	Gly
Glu	Leu	Asn	Lys	Leu	Pro	Phe	Gly	Gly	Val	Trp	Asn	Ala	Val	Leu	Lys	
				165					170					175	Lys	
Leu	Asn	Val	Lys	Arg	Arg	Tyr	Thr	Thr	Tyr	Gly	Thr	Tyr	Thr	Ile	Asn	
				180					185					190	Asn	
Ile	Thr	Val	Asn	Leu	Thr	Asp	Lys	Gly	Asn	Ile	Gln	Ile	Trp	Leu	Pro	
				195					200					205	Pro	
Gln	Phe	Lys	Ser	Asn	Ala	Arg	Val	Asp	Leu	Asn	Leu	Arg	Pro	Thr	Gly	
				210					215					220	Gly	
Gly	Gly	Thr	Tyr	Ile	Gly	Arg	Asn	Ser	Val	Asp	Met	Cys	Phe	Tyr	Asp	
225					230					235					240	Asp
Gly	Tyr	Ser	Thr	Met	Ser	Ser	Ser	Leu	Glu	Ile	Arg	Phe	Gln	Asp	Asp	
				245					250					255	Asp	
Asn	Ser	Lys	Ser	Asp	Gly	Lys	Phe	Tyr	Leu	Lys	Lys	Ile	Asn	Asp	Asp	
				260					265					270	Asp	
Ser	Lys	Glu	Leu	Val	Tyr	Thr	Leu	Ser	Leu	Leu	Leu	Ala	Gly	Lys	Asn	
				275					280					285	Asn	
Leu	Thr	Pro	Thr	Asn	Gly	Gln	Ala	Leu	Asn	Ile	Asn	Thr	Ala	Ser	Leu	
				290					295					300	Leu	
Glu	Thr	Asn	Trp	Asn	Arg	Ile	Thr	Ala	Val	Thr	Met	Pro	Glu	Ile	Ser	
305					310					315					320	Ser
Val	Pro	Val	Leu	Cys	Trp	Pro	Gly	Arg	Leu	Gln	Leu	Asp	Ala	Lys	Val	
				325					330					335	Val	
Lys	Asn	Pro	Glu	Ala	Gly	Gln	Tyr	Met	Gly	Asn	Ile	Lys	Ile	Thr	Phe	
				340					345					350	Phe	
Thr	Pro	Ser	Ser	Gln	Thr	Leu										
				355												

<210> 30

<212> PRT

$\langle 220 \rangle$

<400> 30

Ser Leu Phe Asp Arg Met Thr Phe Leu Cys Met Ser Ser Thr Asp Ala
 65 70 75 80
 Ser Lys Gly Ala Cys Pro Thr Gly Glu Asn Ser Lys Ser Ser Gln Gly
 85 90 95
 Glu Thr Asn Ile Lys Leu Ile Phe Thr Glu Lys Lys Ser Leu Ala Arg
 100 105 110
 Lys Thr Leu Asn Leu Lys Gly Tyr Lys Arg Phe Leu Tyr Glu Ser Asp
 115 120 125
 Arg Cys Ile His Tyr Val Asp Lys Met Asn Leu Asn Ser His Thr Val
 130 135 140
 Lys Cys Val Gly Ser Phe Thr Arg Gly Val Asp Phe Thr Leu Tyr Ile
 145 150 155 160
 Pro Gln Gly Glu Ile Asp Gly Leu Leu Thr Gly Gly Ile Trp Lys Ala
 165 170 175
 Thr Leu Glu Leu Arg Val Lys Arg His Tyr Asp Tyr Asn His Gly Thr
 180 185 190
 Tyr Lys Val Asn Ile Thr Val Asp Leu Thr Asp Lys Gly Asn Ile Gln
 195 200 205
 Val Trp Thr Pro Lys Phe His Ser Asp Pro Arg Ile Asp Leu Asn Leu
 210 215 220
 Arg Pro Glu Gly Asn Gly Lys Tyr Ser Gly Ser Asn Val Leu Glu Met
 225 230 235 240
 Cys Leu Tyr Asp Gly Tyr Ser Thr His Ser Gln Ser Ile Glu Met Arg
 245 250 255
 Phe Gln Asp Asp Ser Gln Thr Gly Asn Asn Glu Tyr Asn Leu Ile Lys
 260 265 270
 Thr Gly Glu Pro Leu Lys Lys Leu Pro Tyr Lys Leu Ser Leu Leu Leu
 275 280 285
 Gly Gly Arg Glu Phe Tyr Pro Asn Asn Gly Lys Ala Phe Thr Ile Asn
 290 295 300
 Asp Thr Ser Ser Leu Phe Ile Asn Trp Asn Arg Ile Lys Ser Val Ser
 305 310 315 320
 Leu Pro Gln Ile Ser Ile Pro Val Leu Cys Trp Pro Ala Asn Leu Thr
 325 330 335
 Phe Met Ser Glu Leu Asn Asn Pro Glu Ala Gly Glu Tyr Ser Gly Ile
 340 345 350
 Leu Asn Val Thr Phe Thr Pro Ser Ser Ser Ser Leu
 355 360

<210> 31

<211> 362

<212> PRT

<213> Artificial Sequence

<220>

<223> ETEC Protein Homology Sequence

<400> 31

Met Lys Lys Ile Phe Ile Phe Leu Ser Ile Ile Phe Ser Ala Val Val
 1 5 10 15
 Ser Ala Gly Arg Tyr Pro Glu Thr Thr Val Gly Asn Leu Thr Lys Ser
 20 25 30
 Phe Gln Ala Pro Arg Leu Asp Arg Ser Val Gln Ser Pro Ile Tyr Asn
 35 40 45

THE UNIVERSITY OF CHICAGO

<211> 353

<213> Artificial Sequence

<223> ETEC Protein Homology Sequence

```
Met Ser Asn Ile Cys Lys Trp Thr Ser Met Thr Ala His Trp Ser Ala
  1                               5               10             15
Ile Ile Asn Phe Ile Arg Lys Tyr Val Tyr Pro Ala Arg Ile Ile Ala
      20                               25             30
```

Ile Leu Ala Gly Ala Thr Leu Pro Gln Val Ala Asp Ala Ile Thr Val
35 40 45
Asp Leu Asn Tyr Asp Lys Asn Asn Val Ala Val Ile Thr Pro Val Trp
50 55 60
Ser Gln Glu Trp Ser Val Ala Asn Val Leu Gly Gly Trp Val Cys Arg
65 70 75 80
Ser Asn Arg Asn Glu Asn Glu Gly Cys Glu Glu Thr His Leu Val Trp
85 90 95
Trp Tyr Ala Phe Gly Ala Tyr Ser Ile Arg Leu Arg Phe Arg Glu Gln
100 105 110
Ile Ser His Ala Glu Ile Thr Leu Ile Leu Leu Gly Ser Val Arg Asp
115 120 125
Ala Cys Thr Gly Val Ile Asn Met Asn Ala Ala Ala Cys Gln Trp Gly
130 135 140
Arg Ser Leu Lys Leu Arg Ile Pro Ser Glu Glu Leu Ala Lys Ile Pro
145 150 155 160
Thr Ser Gly Thr Trp Lys Ala Thr Leu Val Leu Asp Tyr Leu Gln Trp
165 170 175
Gly Gly Asp Asp Pro Leu Gly Thr Ser Thr Thr Asp Ile Thr Leu Asn
180 185 190
Val Thr Asp His Phe Ala Glu Asn Ala Ala Ile Tyr Phe Pro Gln Phe
195 200 205
Gly Thr Ala Thr Pro Arg Val Asp Leu Asn Leu His Arg Met Asn Ala
210 215 220
Ser Gln Met Ser Gly Arg Ala Asn Leu Asp Met Cys Leu Tyr Asp Gly
225 230 235 240
Gly Val Lys Ala Arg Ser Leu Gln Met Met Glu Gly Ser Asn Lys Ser
245 250 255
Gly Thr Gly Phe Gln Val Ile Lys Ser Asp Ser Ala Asp Thr Ile Asp
260 265 270
Tyr Ala Val Ser Met Asn Tyr Gly Gly Arg Ser Ile Pro Val Thr Arg
275 280 285
Gly Val Glu Phe Ser Leu Asp Asn Val Asp Lys Ala Ala Thr Arg Pro
290 295 300
Val Val Leu Pro Gly Gln Arg Gln Ala Val Arg Cys Val Pro Val Pro
305 310 315 320
Leu Thr Leu Thr Thr Gln Pro Phe Asn Ile Arg Glu Lys Arg Ser Gly
325 330 335
Glu Tyr Gln Gly Thr Leu Thr Val Thr Met Leu Met Gly Thr Gln Thr
340 345 350
Pro

<210> 33

<211> 165

<212> PRT

<213> Artificial Sequence

<220>

<223> ETEC Protein Homology Sequence

<400> 33

Met Lys Leu Lys Lys Thr Ile Gly Ala Met Ala Leu Thr Thr Met Phe
1 5 10 15

Val	Ala	Met	Ser	Ala	Ser	Ala	Val	Glu	Lys	Asn	Ile	Thr	Val	Thr	Ala
			20					25					30		
Ser	Val	Asp	Pro	Thr	Ile	Asp	Ile	Leu	Gln	Ala	Asp	Gly	Ser	Ser	Leu
		35					40					45			
Pro	Thr	Ala	Val	Glu	Leu	Thr	Tyr	Ser	Pro	Ala	Ala	Ser	Arg	Phe	Glu
		50				55					60				
Asn	Tyr	Lys	Ile	Ala	Thr	Lys	Val	His	Thr	Asn	Val	Ile	Asn	Lys	Asn
65					70					75				80	
Val	Leu	Val	Lys	Leu	Val	Asn	Asp	Pro	Lys	Leu	Thr	Asn	Val	Leu	Asp
			85						90				95		
Ser	Thr	Lys	Gln	Leu	Pro	Ile	Thr	Val	Ser	Tyr	Gly	Gly	Lys	Leu	Ser
			100					105					110		
Thr	Ala	Asp	Val	Thr	Phe	Glu	Pro	Ala	Glu	Leu	Asn	Phe	Gly	Thr	Ser
		115					120					125			
Gly	Val	Thr	Gly	Val	Ser	Ser	Ser	Gln	Asp	Leu	Val	Ile	Gly	Ala	Thr
		130				135					140				
Thr	Ala	Gln	Ala	Pro	Ser	Ala	Asn	Tyr	Ser	Gly	Val	Val	Ser	Ile	Leu
145					150					155					160
Met	Thr	Leu	Ala	Ser											
				165											

<210> 34

<211> 168

<212> PRT

<213> Artificial Sequence

<220>

<223> ETEC Protein Homology Sequence

<400> 34

Met	Lys	Phe	Lys	Lys	Thr	Ile	Gly	Ala	Met	Ala	Leu	Thr	Thr	Met	Phe
1				5					10					15	
Val	Ala	Val	Ser	Ala	Ser	Ala	Val	Glu	Lys	Asn	Ile	Thr	Val	Thr	Ala
			20					25					30		
Ser	Val	Asp	Pro	Ala	Ile	Asp	Leu	Gln	Ala	Asp	Gly	Asn	Ala	Leu	
		35					40				45				
Pro	Ser	Val	Lys	Leu	Ala	Tyr	Ser	Pro	Ala	Ser	Lys	Ile	Phe	Glu	Ser
		50				55					60				
Tyr	Arg	Val	Met	Thr	Gln	Val	His	Thr	Asn	Asp	Ala	Thr	Lys	Lys	Val
65					70					75				80	
Ile	Val	Lys	Leu	Ala	Asp	Thr	Pro	Gln	Leu	Thr	Asp	Val	Leu	Asn	Ser
			85						90				95		
Thr	Val	Gln	Met	Pro	Ile	Ser	Val	Ser	Trp	Gly	Gly	Val	Leu	Ser	Thr
			100					105					110		
Thr	Ala	Lys	Glu	Phe	Glu	Ala	Ala	Ala	Leu	Gly	Tyr	Ser	Ala	Ser	Gly
		115					120					125			
Val	Asn	Gly	Val	Ser	Ser	Ser	Ser	Gln	Glu	Leu	Val	Ile	Ser	Ala	Pro
		130				135					140				
Lys	Thr	Ala	Gly	Thr	Ala	Pro	Thr	Ala	Gly	Asn	Tyr	Ser	Gly	Val	Val
145					150					155					160
Ser	Leu	Val	Met	Thr	Leu	Gly	Ser								
				165											

<210> 35
 <211> 170
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> ETEC Protein Homology Sequence

<400> 35
 Met Lys Leu Lys Lys Thr Ile Gly Ala Met Ala Leu Ala Thr Leu Phe
 1 5 10 15
 Ala Thr Met Gly Ala Ser Ala Val Glu Lys Thr Ile Ser Val Thr Ala
 20 25 30
 Ser Val Asp Pro Thr Val Asp Leu Leu Gln Ser Asp Gly Ser Ala Leu
 35 40 45
 Pro Asn Val Ala Leu Thr Tyr Ser Pro Ala Val Asn Asn Phe Glu Ala
 50 55 60
 His Thr Ile Asn Thr Val Val His Thr Asn Asp Ser Asp Lys Gly Val
 65 70 75 80
 Val Val Lys Leu Ser Ala Asp Pro Val Leu Ser Asn Val Leu Asn Pro
 85 90 95
 Thr Leu Gln Ile Pro Val Ser Val Asn Phe Ala Gly Lys Pro Leu Ser
 100 105 110
 Thr Thr Gly Ile Thr Ile Asp Ser Asn Asp Leu Asn Phe Ala Ser Ser
 115 120 125
 Gly Val Asn Tyr Val Ser Ser Thr Gln Lys Leu Ser Ile His Ala Asp
 130 135 140
 Ala Thr Arg Val Thr Gly Gly Ala Leu Thr Ala Gly Gln Tyr Gln Gly
 145 150 155 160
 Leu Val Ser Ile Ile Leu Thr Lys Ser Thr
 165 170

<210> 36
 <211> 170
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> ETEC Protein Homology Sequence

<400> 36
 Met Lys Leu Asn Lys Ile Ile Gly Ala Leu Val Leu Ser Ser Thr Phe
 1 5 10 15
 Val Ser Met Gly Ala Ser Ala Ala Glu Lys Asn Ile Thr Val Thr Ala
 20 25 30
 Ser Val Asp Pro Thr Ile Asp Leu Met Gln Ser Asp Gly Thr Ala Leu
 35 40 45
 Pro Ser Ala Val Asn Ile Ala Tyr Leu Pro Gly Glu Lys Arg Phe Glu
 50 55 60
 Ser Ala Arg Ile Asn Thr Gln Val His Thr Asn Asn Lys Thr Lys Gly
 65 70 75 80
 Ile Gln Ile Lys Leu Thr Asn Asp Asn Val Val Met Thr Asn Leu Ser
 85 90 95
 Asp Pro Ser Lys Thr Ile Pro Leu Glu Val Ser Phe Ala Gly Thr Lys

	100		105		110										
Leu	Ser	Thr	Ala	Ala	Thr	Ser	Ile	Thr	Ala	Asp	Gln	Leu	Asn	Phe	Gly
	115						120					125			
Ala	Ala	Gly	Val	Glu	Thr	Val	Ser	Ala	Thr	Lys	Glu	Leu	Val	Ile	Asn
	130						135				140				
Ala	Gly	Ser	Thr	Gln	Gln	Thr	Asn	Ile	Val	Ala	Gly	Asn	Tyr	Gln	Gly
145				150					155						160
Leu	Val	Ser	Ile	Val	Leu	Thr	Gln	Glu	Pro						
			165						170						

<210> 37
 <211> 168
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> ETEC Protein Homology Sequence

	1		5		10		15
Met	Lys	Leu	Lys	Tyr	Thr	Ile	Phe
Val	Ala	Val	Ser	Ala	Ser	Ala	Ala
	20					25	30
Ser	Val	Asp	Pro	Thr	Ile	Asp	Leu
	35					40	45
Pro	Thr	Ala	Val	Asp	Leu	Thr	Tyr
	50					55	60
Asn	Tyr	Ser	Val	Leu	Thr	Gln	Ile
65				70			75
Leu	Asp	Val	Arg	Leu	Val	Asp	Thr
			85				90
Pro	Thr	Ser	Thr	Ile	Pro	Leu	Thr
	100					105	110
Ser	Thr	Ser	Ala	Gln	Lys	Ile	Ala
	115					120	125
Thr	Gly	Thr	Ala	Gly	Val	Ser	Asn
	130					135	140
Thr	Thr	Ser	Gly	Lys	Pro	Ser	Ala
145				150			155
Ile	Val	Met	Thr	Gln	Ser	Thr	Asn
				165			

<210> 38
 <211> 142
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> ETEC Protein Homology Sequence

	1		5		10		15
Val	Asp	Pro	Thr	Ile	Asp	Ile	Leu
							Gln
							Ala
							Asn
							Gly
							Ser
							Ala
							Leu
							Pro

Thr Ala Val Asp Leu Thr Tyr Leu Pro Gly Ala Lys Thr Phe Glu Asn
 20 25 30
 Tyr Ser Val Leu Thr Gln Ile Tyr Thr Asn Asp Pro Ser Lys Gly Leu
 35 40 45
 Asp Val Arg Leu Val Asp Thr Pro Lys Leu Thr Asn Ile Leu Gln Pro
 50 55 60
 Thr Ser Thr Ile Pro Leu Thr Val Ser Trp Ala Gly Lys Thr Leu Ser
 65 70 75 80
 Thr Ser Ala Gln Lys Ile Ala Val Gly Asp Leu Gly Phe Gly Ser Thr
 85 90 95
 Gly Thr Ala Gly Val Ser Asn Ser Lys Glu Leu Val Ile Gly Ala Thr
 100 105 110
 Thr Ser Gly Thr Ala Pro Ser Ala Gly Lys Tyr Gln Gly Val Val Ser
 115 120 125
 Ile Val Met Thr Gln Ser Thr Asp Thr Ala Ala Pro Val Pro
 130 135 140

<210> 39

<211> 133

<212> PRT

<213> Artificial Sequence

<220>

<223> ETEC Protein Homology Sequence

<400> 39

Val Asp Pro Lys Leu Asp Leu Leu Gln Ala Asp Gly Thr Ser Leu Pro
 1 5 10 15
 Asp Ser Ile Ala Leu Thr Tyr Ser Ser Ala Ser Asn Asn Phe Glu Val
 20 25 30
 Tyr Ser Leu Asn Thr Ala Ile His Thr Asn Asp Lys Thr Lys Ala Val
 35 40 45
 Val Val Lys Leu Ser Ala Pro Ala Val Leu Ser Asn Ile Met Lys Pro
 50 55 60
 Ser Ser Gln Ile Pro Met Lys Val Thr Leu Gly Gly Lys Thr Leu Ser
 65 70 75 80
 Thr Ala Asp Ala Glu Phe Ala Ala Asp Thr Leu Asn Phe Gly Ala Ser
 85 90 95
 Gly Val Glu Asn Val Ser Ser Val Gln Leu Thr Ile His Ala Glu
 100 105 110
 Ala Ala Pro Pro Glu Ala Gly Asn Tyr Gln Gly Val Ile Ser Leu Ile
 115 120 125
 Met Thr Gln Lys Thr
 130

<210> 40

<211> 134

<212> PRT

<213> Artificial Sequence

<220>

<223> ETEC Protein Homology Sequence

<400> 40

Val Asp Pro Lys Leu Asp Leu Leu Gln Ala Asp Gly Thr Ser Leu Pro
1 5 10 15
Asp Ser Ile Ala Leu Thr Tyr Ser Ser Ala Ser Asn Asn Phe Glu Val
20 25 30
Tyr Ser Leu Asn Thr Ala Ile His Thr Asn Asp Lys Ser Lys Gly Val
35 40 45
Val Val Lys Leu Ser Ala Ser Pro Val Leu Ser Asn Ile Met Pro Asn
50 55 60
Ser Gln Ile Pro Met Lys Val Thr Leu Gly Gly Glu Thr Leu Asn Thr
65 70 75 80
Thr Asp Thr Glu Phe Thr Val Asp Thr Leu Asn Phe Gly Thr Ser Gly
85 90 95
Val Glu Asn Val Ser Ser Thr Gln Gln Leu Thr Ile His Ala Asp Thr
100 105 110
Gln Gly Thr Ala Pro Glu Ala Gly Asn Tyr Gln Gly Ile Ile Ser Leu
115 120 125
Ile Met Thr Gln Lys Thr
130

100240-1636360